

Cognitive Limits on Rationality

IN THE THREE PREVIOUS CHAPTERS, we have considered how motivations and goals affect human behavior in organizations. The content of these chapters constitutes an important amendment to the "classical" theory of organization, which regards the employee as an "instrument." In the present chapter and the following one we shall focus on a different set of qualities of the organization member—his characteristics as a rational man. When, at the end of Chapter 7, we conclude our study of these characteristics and their implications for organization theory, we will have completed the main tasks we set ourselves:

1. to eliminate, one by one, the artificialities of the classical description of the employee as an instrument;
2. to replace this abstraction with a new one that recognizes that members of organizations have wants, motives, and drives, and are limited in their knowledge and in their capacities to learn and to solve problems.

First, we take note of some characteristics of human rationality that bear upon decision-making processes in organizations. Next, we see how organizational decision-making is organized into "programs" or strategies. In the third section we re-examine the phenomenon of organizational identification in the light of the introductory analysis to see to what extent identification is an intellective rather than a motivational process. The fourth section considers the implications of the decision-making process for the division of work; and the fifth section discusses the communications requirements and processes that arise out of the division of work. In the final section we state some of the broader propositions about organization structure that can be derived from the analysis of decision-making processes.

6.1. THE CONCEPT OF RATIONALITY

How does the rationality of "administrative man" compare with that of classical "economic man" or with the rational man of modern statistical decision theory? The rational man of economics and statistical decision theory makes "optimal" choices in a highly specified and clearly defined environment:

1. When we first encounter him in the decision-making situation, he already has laid out before him the whole set of alternatives from which he will choose his action. This set of alternatives is simply "given"; the theory does not tell how it is obtained.
2. To each alternative is attached a set of consequences—the events that will ensue if that particular alternative is chosen. Here the existing theories fall into three categories: (a) *Certainty*: theories that assume the decision maker has complete and accurate knowledge of the consequences that will follow on each alternative. (b) *Risk*: theories that assume accurate knowledge of a probability distribution of the consequences of each alternative. (c) *Uncertainty*: theories that assume that the consequences of each alternative belong to some subset of all possible consequences, but that the decision maker cannot assign definite probabilities to the occurrence of particular consequences.

3. At the outset, the decision maker has a "utility function" or a "preference-ordering" that ranks all sets of consequences from the most preferred to the least preferred.

4. The decision maker selects the alternative leading to the preferred set of consequences. In the case of *certainty*, the choice is unambiguous. In the case of *risk*, rationality is usually defined as the

choice of that alternative for which the expected utility is greatest. Expected utility is defined here as the average, weighted by the probabilities of occurrence, of the utilities attached to all possible consequences. In the case of *uncertainty*, the definition of rationality becomes problematic. One proposal that has had wide currency is the rule of "minimax risk": consider the worst set of consequences that may follow from each alternative, then select the alternative whose "worst set of consequences" is preferred to the worst sets attached to other alternatives. There are other proposals (e.g., the rule of "minimax regret"), but we shall not discuss them here.

Some difficulties in the classical theory. There are difficulties with this model of rational man. In the first place, only in the case of certainty does it agree well with common-sense notions of rationality. In the case of uncertainty, especially, there is little agreement, even among exponents of statistical decision theory, as to the "correct" definition, or whether, indeed, the term "correct" has any meaning here (Marschak, 1950).

A second difficulty with existing models of rational man is that it makes three exceedingly important demands upon the choice-making mechanism. It assumes (1) that all the alternatives of choice are "given"; (2) that all the consequences attached to each alternative are known (in one of the three senses corresponding to certainty, risk, and uncertainty respectively); (3) that the rational man has a complete utility-ordering (or cardinal function) for all possible sets of consequences.

One can hardly take exception to these requirements in a normative model—a model that tells people how they *ought* to choose. For if the rational man lacked information, he might have chosen differently "if only he had known." At best, he is "subjectively" rational, not "objectively" rational. But the notion of objective rationality assumes there is some objective reality in which the "real" alternatives, the "real" consequences, and the "real" utilities exist. If this is so, it is not even clear why the cases of choice under risk and under uncertainty are admitted as rational. If it is not so, it is not clear why only limitations upon knowledge of consequences are considered, and why limitations upon knowledge of alternatives and utilities are ignored in the model of rationality.

From a phenomenological viewpoint we can only speak of rationality relative to a frame of reference; and this frame of reference will be determined by the limitations on the rational man's knowledge. We can, of course, introduce the notion of a person observing the

choices of a subject, and can speak of the rationality of the subject relative to the frame of reference of the observer. If the subject is a rat and the observer is a man (especially if he is the man who designed the experimental situation), we may regard the man's perception of the situation as objective and the rat's as subjective. (We leave out of account the specific difficulty that the rat presumably knows his own utility function better than the man does.) If, however, both subject and observer are men—and particularly if the situation is a natural one not constructed for experimental purposes by the observer—then it becomes difficult to specify the objective situation. It will be safest, in such situations, to speak of rationality only relative to some specified frame of reference.

The classical organization theory described in Chapter 2, like classical economic theory, failed to make explicit this subjective and relative character of rationality, and in so doing, failed to examine some of its own crucial premises. The organizational and social environment in which the decision maker finds himself determines what consequences he will anticipate, what ones he will not; what alternatives he will consider, what ones he will ignore. In a theory of organization these variables cannot be treated as unexplained independent factors, but must themselves be determined and predicted by the theory.

Routinized and problem-solving responses. The theory of rational choice put forth here incorporates two fundamental characteristics: (1) Choice is always exercised with respect to a limited, approximate, simplified "model" of the real situation [A-6.1]. We call the chooser's model his "definition of the situation." (2) The elements of the definition of the situation are not "given"—that is, we do not take these as data of our theory—but are themselves the outcome of psychological and sociological processes, including the chooser's own activities and the activities of others in his environment [A-6.2] (Simon, 1947, 1955; March, 1955a; Cyert and March, 1955, 1956; Newell, Shaw, and Simon, 1958).

Activity (individual or organizational) can usually be traced back to an environmental stimulus of some sort, e.g., a customer order or a fire gong. The responses to stimuli are of various kinds. At one extreme, a stimulus evokes a response—sometimes very elaborate—that has been developed and learned at some previous time as an appropriate response for a stimulus of this class. This is the "routinized" end of the continuum, where a stimulus calls forth a performance program almost instantaneously.

At the other extreme, a stimulus evokes a larger or smaller amount of problem-solving activity directed toward finding performance ac-

tivities with which to complete the response. Such activity is distinguished by the fact that it can be dispensed with once the performance program has been learned. Problem-solving activities can generally be identified by the extent to which they involve *search*: search aimed at discovering alternatives of action or consequences of action. "Discovering" alternatives may involve inventing and elaborating whole performance programs where these are not already available in the problem solver's repertory (Katona, 1951).

When a stimulus is of a kind that has been experienced repeatedly in the past, the response will ordinarily be highly routinized [A-6.3]. The stimulus will evoke, with a minimum of problem-solving or other computational activity, a well-structured definition of the situation that will include a repertory of response programs, and programs for selecting an appropriate specific response from the repertory. When a stimulus is relatively novel, it will evoke problem-solving activity aimed initially at constructing a definition of the situation and then at developing one or more appropriate performance programs [A-6.4].

Psychologists (e.g., Wertheimer, Duncker, de Groot, Maier) and observant laymen (e.g., Poincaré, Hadamard) who have studied creative thinking and problem-solving have been unanimous in ascribing a large role in these phenomena to search processes. Search is partly random, but in effective problem-solving it is not blind. The design of the search process is itself often an object of rational decision. Thus, we may distinguish substantive planning—developing new performance programs—from procedural planning—developing programs for the problem-solving process itself. The response to a particular stimulus may involve more than performance—the stimulus may evoke a spate of problem-solving activity—but the problem-solving activity may itself be routinized to a greater or lesser degree. For example, search processes may be systematized by the use of check lists.

Satisfactory versus optimal standards. What kinds of search and other problem-solving activity are needed to discover an adequate range of alternatives and consequences for choice depends on the criterion applied to the choice. In particular, finding the optimal alternative is a radically different problem from finding a satisfactory alternative. An alternative is *optimal* if: (1) there exists a set of criteria that permits all alternatives to be compared, and (2) the alternative in question is preferred, by these criteria, to all other alternatives. An alternative is *satisfactory* if: (1) there exists a set of criteria that describes minimally satisfactory alternatives, and (2) the alternative in question meets or exceeds all these criteria.

Most human decision-making, whether individual or organizational,

is concerned with the discovery and selection of satisfactory alternatives; only in exceptional cases is it concerned with the discovery and selection of optimal alternatives [A-6.5]. To optimize requires processes several orders of magnitude more complex than those required to satisfice. An example is the difference between searching a haystack to find the *sharpest* needle in it and searching the haystack to find a needle sharp enough to sew with.

In making choices that meet satisfactory standards, the standards themselves are part of the definition of the situation. Hence, we need not regard these as given—any more than the other elements of the definition of the situation—but may include in the theory the processes through which these standards are set and modified. The standard-setting process may itself meet standards of rationality: for example, an "optimizing" rule would be to set the standard at the level where the marginal improvement in alternatives obtainable by raising it would be just balanced by the marginal cost of searching for alternatives meeting the higher standard. Of course, in practice the "marginal improvement" and the "marginal cost" are seldom measured in comparable units, or with much accuracy. Nevertheless, a similar result would be automatically attained if the standards were raised whenever alternatives proved easy to discover, and lowered whenever they were difficult to discover. Under these circumstances, the alternatives chosen would not be far from the optima, if the cost of search were taken into consideration. Since human standards tend to have this characteristic under many conditions, some theorists have sought to maintain the optimizing model by introducing cost-of-search considerations. Although we doubt whether this will be a fruitful alternative to the model we are proposing in very many situations, neither model has been used for predictive purposes often enough to allow a final judgment.

Performance programs. We have seen that under certain circumstances the search and choice processes are very much abridged. At the limit, an environmental stimulus may evoke immediately from the organization a highly complex and organized set of responses. Such a set of responses we call a *performance program*, or simply a *program*. For example, the sounding of the alarm gong in a fire station initiates such a program. So does the appearance of a relief applicant at a social worker's desk. So does the appearance of an automobile chassis in front of the work station of a worker on the assembly line.

Situations in which a relatively simple stimulus sets off an elaborate program of activity without any apparent interval of search, problem-solving, or choice are not rare. They account for a very large part of

the behavior of all persons, and for almost all of the behavior of persons in relatively routine positions. Most behavior, and particularly most behavior in organizations, is governed by performance programs.

The term "program" is not intended to connote complete rigidity. The content of the program may be adaptive to a large number of characteristics of the stimulus that initiates it. Even in the simple case of the fire gong, the response depends on the location of the alarm, as indicated by the number of strokes. The program may also be conditional on data that are independent of the initiating stimuli.

It is then more properly called a *performance strategy*. For example, when inventory records show that the quantity on hand of a commodity has decreased to the point where it should be reordered, the decision rule that governs the behavior of the purchasing agent may call upon him to determine the amount to be ordered on the basis of a formula into which he inserts the quantity that has been sold over the past 12 months. In this case, search has been eliminated from the problem, but choice—of a very routinized kind, to be sure—remains.

We will regard a set of activities as routinized, then, to the degree that choice has been simplified by the development of a fixed response to defined stimuli. If search has been eliminated, but a choice remains in the form of a clearly defined and systematic computing routine, we will still say that the activities are routinized. We will regard activities as unroutinized to the extent that they have to be preceded by program-developing activities of a problem-solving kind.

6.2 PERFORMANCE PROGRAMS IN ORGANIZATIONS

There are several ways to determine what programs a particular organization uses:

1. Observing the behavior of organization members. In relatively routine positions, where the same situations recur repetitively and are handled in terms of fairly definite programs, it is easy to infer the program from behavior. This is a common method for inducting new members of an organization into its procedures.

2. Interviewing members of the organization. Most programs are stored in the minds of the employees who carry them out, or in the minds of their superiors, subordinates, or associates. For many purposes, the simplest and most accurate way to discover what a person does is to ask him.

3. Examining documents that describe standard operating procedures. Programs may be written down, more or less completely and

more or less accurately. The relation of a written operating procedure to the actual program that is carried out is complex, for the program may have been written down: (a) as an instruction to initiate a new program and communicate it to those who will carry it out; (b) as a description of an existing program to instruct new organization members; or (c) as an exposition (with or without amendments) of an existing program to legitimize or "formalize" it. There are other possibilities besides these three. In any event, when a document is used as a source of information about a program, the purposes for which it was prepared are relevant to its interpretation.

A person who has been trained in the observation of organizations can extract by these and other techniques a large part of the program that governs routine behavior. This is such a common-sense fact that its importance has been overlooked: Knowledge of the program of an organization permits one to predict in considerable detail the behavior of members of the organization. And the greater the *programming* (6.1) of individual activities in the organization, the greater the *predictability* (6.2) of those activities [6.2:6.1].

To be sure, prediction of behavior from the knowledge of a program has none of the element of "surprise" that we commonly associate with scientific prediction—any more than prediction of the lines that will be uttered by a Hamlet on the stage. It is no less important for its common-sense obviousness.

In general, we would anticipate that programs will be generated by past experience and in expectation of future experience in a given situation. Thus, the greater the *repetitiveness* (6.3) of individual activities, the greater the programming [6.1:6.3]. From this one would predict that programming will be most complete for clerical and factory jobs, particularly when the work is organized largely by process.

The prediction of behavior from a program when tasks are relatively simple and routine is illustrated by findings of Guetzkow and Simon (1955) using five-man experimental groups in the Bavelas network. Employing methods-analysis techniques, they were able to predict average trial times of groups to within 10% from a knowledge of the methods the groups were using to perform the task.

If the program determines in some detail the behavior of individuals and groups performing relatively routine tasks, then we can predict behavior to the extent that we can answer the following questions:

(1) What motivates members of the organization to accept a program as a determinant of their behavior? What processes, other than mo-

tivation, are involved in implementation of programs? This question has already been examined in earlier chapters. (2) What determines the content of a program? To what extent can the program be predicted uniquely from the requirements of the task? How are programs invented and developed, and what are the determinants of this process? (3) What are the consequences of programs, as developed and executed, for the goal and subgoal structure of the organization? (4) What are the predictors of behavior in areas that are not routinized and are unprogrammed? This question will be taken up in the next chapter. †

We turn now to the second and third of these questions.

Program content. The extent to which many human activities, both manual and clerical, can be programmed is shown by the continuing spread of automation to encompass a wider and wider range of tasks. In order to substitute automatic processes for human operatives, it is necessary to describe the task in minute detail, and to provide for the performance of each step in it. The decomposition of tasks into their elementary program steps is most spectacularly illustrated in modern computing machines which may carry out programs involving thousands of such steps. The capabilities of computers have now been extended to many tasks that until recently have been thought to be relatively complex, involving problem-solving activities of a fairly high order. Some examples are several existing computer programs for the automatic design of small electric motors and transformers, a program that enables a computer to discover proofs for certain kinds of mathematical theorems, and a program for translating languages.

Even on routine jobs, *program content* (6.4) varies. We have already mentioned the extreme case: the detailed specification of output, methods, and pace in a man-paced assembly operation. But not all programs are of this type. They may not contain detailed time specifications (e.g., in typical machine-paced operations). In fact, programs usually specify the content of an activity more closely than its timing [A-6.6]. They may specify the properties of the product (e.g., in blueprints, tolerances, etc.) rather than the detail of the methods to be used. We need propositions that will explain variations in program content along these dimensions:

- (a) The extent to which pacing rules are built into the program.
- (b) The extent to which work activities are detailed in the program.
- (c) The extent to which product specifications are detailed in the program.

Since performance programs are important aspects of the organizational system, their content will presumably tend to be related to the

functions they perform. We can identify two major functions that such programs fulfill, or at least are intended to fulfill. First, they are a part of the control system in the organization. Organizations attempt to control employees by specifying a standard operating procedure and attaching organizational rewards and penalties to it. Second, performance programs are important parts of the coordination system in the organization. They help fulfill the needs for interdepartmental predictability [A-6.7] (Blau, 1955).

Insofar as they are to function as controls, the programs must be linked to variables that are observable and measurable. We would expect program content to be a function of the *ease of observing job activities* (6.5), the *ease of observing job output* (6.6), and the *ease of relating activities to output* (6.7) [6.4:6.5, 6.6, 6.7]. Thus, we would predict that programs will contain activity specifications in preference to product specifications to the extent that: (a) the activity pattern is easily observed and supervised; (b) the quantity and quality of output are not easily observed and supervised; (c) the relations between activity pattern and output are highly technical, and are matters of scientific and engineering knowledge, better known to specialists in the organization than to the operatives (Ridley and Simon, 1938).

Conversely, programs will contain specifications of quality and quantity of output to the extent that: (a) the activity pattern is difficult to observe and supervise; (b) the quantity and quality of output are easily observed and supervised; (c) the relations between activity pattern and output are matters of common sense, are matters of skill in the specific occupation for which the operatives are trained, or are highly variable, depending upon circumstances of the individual situation that are better known to the operatives than to supervisors and specialists.

For performance programs to serve as coordinative devices, they must be linked to the coordination needs that are felt by the organization. Consequently, we would hypothesize that program content will be a function of the *need for activity coordination* (6.8) and the *need for output coordination* (6.9) [6.4:6.8, 6.9]. The more minutely other members of the organization need to synchronize or coordinate their activities with the activities of a particular member, the more completely will the program specify the activity pattern and/or the pacing of those activities. But to the extent that the activities of the former depend on the characteristics of the output of the latter, rather than on his activities, the program will specify product characteristics.

These propositions about program content are derived from the assumption that the program will be rationally adapted to the organiza-

tion's objectives. To the extent that this assumption actually determines program, program content becomes a technological question in exactly the same way as the form of the production function is a technological question. In the experiment with the Bavelas network, mentioned previously, determining the most efficient program for performing the task is an exercise in methods study resting upon knowledge of human physiological constants—the times required to perform certain simple acts. If we assume that over some period of time an organization will actually arrive at an efficient program, we can predict its long-run behavior from our technical analysis.

Suppose, however, that we substitute for the maximizing assumption implicit in this method of prediction the assumption that behavior is rational in the more limited sense described earlier: that programs are sought that will operate "satisfactorily," and that the "best" program is not necessarily sought or found. In this case, predicting the program becomes more difficult. Which of the (presumably numerous) satisfactory potential programs the organization will adopt depends, under these circumstances, upon the procedures it employs to construct new programs and to improve existing ones. These procedures will provide the principal subject matter for the next chapter.

The structure of programs. To illustrate further the structure of programs for handling recurrent events, we will describe some formal procedures often used by business concerns for controlling inventory. We will analyze first the common "two-bin" system of inventory control, then a more elaborate system.

In the two-bin system of inventory control, two quantities are established for each item kept in stock: (1) the order quantity (the amount to be purchased on a single order), (2) the buffer stock (the amount that should be on hand when a new order is placed). The program is very simple:

"1. When material is drawn from stock, note whether the quantity that remains equals or exceeds the buffer stock. If not:

"2. Write a purchase order for the specified order quantity."

Let us call the first step the "program-evoking" step, and the second step the "program-execution" step. The bifurcation is characteristic of programs—a program includes a specification of the circumstances under which the program is to be evoked [A-6.8]. In the example just cited, the program specifies certain observations, which are to be made (whether the buffer stock is intact) whenever a certain event occurs (withdrawal of material from stock). A decision to act or not

to act (to apply or not to apply the program) is based on the result of the observation.

The program-evoking step may involve only observation auxiliary to some other activity (as in this example), or it may invoke systematic scanning of some part of the environment (e.g., the activity of a quality inspector). Further, a program-execution step by one member of an organization may serve as a program-evoking step for another member. In the example above, the receipt of a purchase order from the inventory clerk is a program-evoking step for the purchasing department.

In our very simple example, the program-execution step requires neither discretion nor problem-solving. In more complicated situations, the program will be a strategy; i.e., action will be contingent on various characteristics of the situation. For example, in a more elaborate inventory control scheme, the purchase quantity may depend on a forecast of sales. Then the program might look like this:

"1. When material is drawn from stock, note whether the quantity that remains equals or exceeds the buffer stock. If not:

"2. Determine from the sales forecast provided by the sales department the sales expected in the next k months.

"3. Insert this quantity in the 'order quantity formula,' and write a purchase order for the quantity thus determined."

This program, although it is contingent on certain changing facts (the sales forecast), does not allow discretion to the person who executes it—at least in ordinary meanings of the word "discretion." If, however, the organization does not provide the inventory clerk with an official sales forecast, or does not establish a specific order quantity, we would say that the clerk's activity was, to that extent, discretionary. We might discover by observation and interview that the clerk was in fact following a very definite and invariable program, but one stored in his own memory and not recorded in official instructions.

The nature of discretion. The amounts and kinds of *discretion* (6.10) available to the organizational participant are a function of his performance program and in particular the extent to which the program specifies activities (means) and the extent to which it specifies product or outcome (ends) [6.10:6.4]. The further the program goes in the latter direction, the more discretion it allows for the person implementing the program to supply the means-end connections. Compare the programs cited earlier with the following alternative program:

"1. It is the duty of the inventory clerk to determine when each item should be recorded and in what quantity, and to place orders with the purchasing department. He should perform this function with attention to the costs of holding inventories, the costs of shortages, and the economies associated with bulk orders."

If we interpret the last sentence as enjoining the clerk to minimize the sum of the specified costs, we see that this program specifies a goal, but leaves the means undetermined. To construct a "rational" program starting from these premises requires the following steps: (1) defining the total cost function in specific terms; (2) estimating the coefficients that appear in the cost function; (3) deriving a formula or "strategy" that specifies the ordering rules as functions of: (a) the coefficients that appear in the cost function, (b) the sales forecasts (i.e., finding the policy that minimizes step 1), and (4) inserting in the formula the coefficients estimated in step 2, and the sales forecasts.

It is difficult to find a place for discretion within the framework of traditional theories of rational behavior. In the present theory, however, a whole host of phenomena fall under this heading.

First, when a program involves search activities, the actual course of action depends on what is found. We may regard the choice of a course of action after search as discretionary.

Second, when a program describes a strategy, application of the strategy to specific circumstances requires forecasts or other estimates of data. We may regard the application of the strategy to select a course of action as discretionary.

Third, a program may exist in the memory of the individual who is to apply it, having arrived there either as a result of extraorganizational training (e.g., professional training or apprenticeship), or as a product of learning from experience rather than as a result of formal instructions. Under these circumstances we often regard him as behaving in a discretionary fashion.

In all of the cases listed above, the decision process may in fact be highly routinized—the term "discretionary" referring in these instances to the form of the performance program or the source from which it was acquired. These cases need to be distinguished from a fourth meaning of "discretionary": A program may specify only general goals, and leave unspecified the exact activities to be used in reaching them. Moreover, knowledge of the means-ends connections may be sufficiently incomplete and inexact that these cannot be very well specified in advance. Then "discretion" refers to the development and modification of the performance program through problem-solving and

learning processes. Although it is difficult to draw a perfectly sharp line between changing a program and changing a datum in applying a strategy, we have already argued that there is an important difference of degree here. With these several meanings of the term "discretionary" in mind, we do not need separate propositions about the amount of discretion, for these will be subsumed under the propositions already noted that specify the form, content, and completeness of programs.

Interrelation of programs. A program, whether simple or complex, is initiated when it is evoked by some stimulus. The whole pattern of programmed activity in an organization is a complicated mosaic of program executions, each initiated by its appropriate program-evoking step [A-6.9].

Insofar as the stimuli that evoke programs come from outside the organization, the individual pieces of this mosaic are related to each other only in making claims on the same time and resources, and hence in posing an allocation problem. Nevertheless, if the goal of optimizing is taken seriously, this allocation problem will usually complicate the problem-solving process greatly, for it requires the marginal return from activity in response to any particular stimulus to be equated with the marginal return from activities in response to all other stimuli. Hence, all programs must be determined simultaneously.

When the goal is to respond to stimuli in a satisfactory, but not necessarily optimal, fashion, choice is much simpler; for the standards may be set at levels that permit a satisficing response to each stimulus without concern for the others. The organization, under these circumstances, normally has some slack that reduces the interdependence among its several performance programs.

Apart from resource-sharing, there may be other and more integral connections among programs. Program A may be a *higher-level* program, i.e., a problem-solving activity whose goal is to revise other programs, either by constructing new ones, reconstructing existing ones, or simply modifying individual premises in existing programs. In this case, the *content* of the lower-level programs that are related to A will depend on A. Or, program A may be a program one of whose execution steps serves as an initiating stimulus for program B.

The inventory example illustrates both possibilities. As to the first, program A may be a forecasting program, or a program for periodic revision of the coefficients in the cost function. As to the second possibility, the order that goes from the inventory clerk to the purchasing department serves to initiate one of the purchasing programs of the latter.

Program and organization structure. In organizations there generally is a considerable degree of parallelism between the hierarchical relations among members of the organization and the hierarchical relations among program elements. That is to say, the programs of members of higher levels of the organization have as their main output the modification or initiation of programs for individuals at lower levels [A-6.10].

Any organization possesses a repertory of programs that, collectively, can deal in a goal-oriented way with a range of situations. As new situations arise, the construction of an entirely new program from detailed elements is rarely contemplated. In most cases, adaptation takes place through a recombination of lower-level programs that are already in existence [A-6.11]. An important objective of standardization is to widen as far as possible the range of situations that can be handled by combination and recombination of a relatively small number of elementary programs.

Limitation of high-level action to the recombination of programs, rather than the detailed construction of new programs out of small elements, is extremely important from a cognitive standpoint. Our treatment of rational behavior rests on the proposition that the "real" situation is almost always far too complex to be handled in detail. As we move upwards in the supervisory and executive hierarchy, the range of interrelated matters over which an individual has purview becomes larger and larger, more and more complex. The growing complexity of the problem can only be matched against the finite powers of the individual if the problem is dealt with in grosser and more aggregative form. One way in which this is accomplished is by limiting the alternatives of action that are considered to the recombination of a repertory of programs (Simon, 1953b).

We may again illustrate this point with the inventory example. Top management decides upon the total dollar inventories without controlling the distribution of inventories among individual items. Specific inventory control programs are found at lower levels of the organization.

6.3 PERCEPTION AND IDENTIFICATIONS

We have seen that humans, whether inside or outside administrative organizations, behave rationally, if at all, only relative to some set of "given" characteristics of the situation. These "givens" include knowledge or assumptions about future events or probability distributions

of future events, knowledge of alternatives available for action, knowledge of consequences attached to alternatives—knowledge that may be more or less complete—and rules or principles for ordering consequences or alternatives according to preference.

These four sets of givens define the situation as it appears to the rational actor. In predicting his behavior, we need this specification and not merely a specification of the situation as it "really" is, or, more precisely, as it appears to an outside observer.

The steps that lead, for an actor, to his defining the situation in a particular way involve a complex interweaving of affective and cognitive processes. What a person wants and likes influences what he sees; what he sees influences what he wants and likes.

In the three previous chapters we have examined primarily motivational and affective factors. We have considered the relation between individual goals and organizational goals, the ways in which goals are acquired from reference groups, and the motivational bases for conformity with group goals. Cognition enters into the definition of the situation in connection with goal attainment—determining what means will reach desired ends. But cognition enters into the goal-formation process also, because the goals used as criteria for choice seldom represent "final" or "ultimate" values. Instead, they too reflect the perceived relations of means to ends and hence are modified by changing beliefs about these relations. Since goals provide the principal bridge between motivations and cognition, we will begin our consideration of cognitive elements in the definition of the situation with the topic of subgoal formation.

Cognitive aspects of subgoal formation. An individual can attend to only a limited number of things at a time. The basic reason why the actor's definition of the situation differs greatly from the objective situation is that the latter is far too complex to be handled in all its detail. Rational behavior involves substituting for the complex reality a model of reality that is sufficiently simple to be handled by problem-solving processes.

In organizations where various aspects of the whole complex problem are being handled by different individuals and different groups of individuals, a fundamental technique for simplifying the problem is to factor it into a number of nearly independent parts, so that each organizational unit handles one of these parts and can omit the others from its definition of the situation [A-6.12]. This technique is also prominent in individual and small-group behavior. A large complex task is broken down into a sequence of smaller tasks, the conjunction of which adds up to the accomplishment of the larger. The factori-

zation of a large task into parts can be more elaborate for an organization than for an individual, but the underlying reason is the same: the definition of the situation at any one moment must be sufficiently simple to be encompassed by a human mind.

The principal way to factor a problem is to construct a means-end analysis. The means that are specified in this way become subgoals which may be assigned to individual organizational units [A-6.13]. This kind of jurisdictional assignment is often called "organization by purpose" or "departmentalization by purpose."

The motivational aspect of this particular process of subgoal formation is rather simple. Whatever will motivate individuals and groups to accept the tasks assigned them through the legitimate (formal and informal) processes of the organization will provide motivation for subgoals. For the subgoals are implicit or explicit in the definition of the situation as it is incorporated in the task assignment.

When tasks have been allocated to an organizational unit in terms of a subgoal, other subgoals and other aspects of the goals of the larger organization tend to be ignored in the decisions of the subunit. In part, this bias in decision-making can be attributed to shifts in the *focus of attention* (6.11). The definition of the situation that the sub-unit employs is simplified by omitting some criteria and paying particular attention to others. In particular, we expect the focus of attention to be a function of the *differentiation of subgoals* (6.12) and the *persistence of subgoals* (6.13) [6.11:6.12, 6.13].

The tendency of members of an organizational unit to evaluate action only in terms of subgoals, even when these are in conflict with the goals of the larger organization, is reinforced by at least three cognitive mechanisms. The first of these is located within the individual decision maker, the second within the organizational unit, and the third in the environment of the organizational unit.

In the individual there is reinforcement through selective perception and rationalization. That is, the persistence of subgoals is furthered by the focus of attention it helps to generate [6.13:6.11]. The propensity of individuals to see those things that are consistent with their established frame of reference is well established in individual psychology. Perceptions that are discordant with the frame of reference are filtered out before they reach consciousness, or are reinterpreted or "rationalized" so as to remove the discrepancy. The frame of reference serves just as much to validate perceptions as the perceptions do to validate the frame of reference.

Within the organizational unit there is reinforcement through the *content of in-group communication* (6.14). Such communication

affects the *focus of information* (6.15) [6.15:6.14], and thereby increases subgoal persistence [6.13:6.15]. The vast bulk of our knowledge of fact is not gained through direct perception but through the second-hand, third-hand, and *n*th-hand reports of the perceptions of others, transmitted through the channels of social communication. Since these perceptions have already been filtered by one or more communicators, most of whom have frames of reference similar to our own, the reports are generally consonant with the filtered reports of our own perceptions, and serve to reinforce the latter. In organizations, two principal types of in-groups are of significance in filtering: in-groups with members in a particular organizational unit, and in-groups with members in a common profession [A-6.14]. Hence, we may distinguish *organizational identifications* and *professional identifications*. There are others, of course, but empirically these appear to be the most significant.

Finally, there is reinforcement through selective exposure to environmental stimuli. The *division of labor in the organization* (6.16) affects the information that various members receive [6.15:6.16]. This differentiation of information contributes to the differentiation of subgoals [6.12:6.15]. Thus perceptions of the environment are biased even before they experience the filtering action of the frame of reference of the perceiver. Salesmen live in an environment of customers; company treasurers in an environment of bankers; each sees a quite distinct part of the world (Dearborn and Simon, 1958).

There is one important distinction between this source of reinforcement and the two mentioned previously. Reinforcement through selective perception and rationalization and reinforcement through in-group communication serve to explain how a particular definition of the situation, once it becomes established in an individual or group, maintains itself with great stability and tenacity. These mechanisms do not explain, however, what particular definitions of the situation will *become* established in particular environments—they explain behavior persistence and not the origins of behavior. In order to predict what particular subgoals we are likely to find in particular parts of an organization, we must take as our starting point (a) the system of subgoal assignment that has resulted from analysis of the organization's goals, and (b) the kinds of stimuli to which each organizational unit is exposed in carrying out its assignments. Under the last heading we must include the selective feedback to organizational units of those consequences of action that relate to their particular subgoals.

Through these mechanisms of subgoal formation and subgoal perception, there is selective attention to particular consequences of pro-

posed alternatives, and selective inattention to others. The magnitude of these effects depends in part on variations in the "capacity" of the individual participants in the organization. The smaller the *span of attention* (6.17), the narrower the focus of attention and the more critical the screening mechanisms cited above [6.11:6.17]. One variable of particular importance in determining the span of attention is, of course, the *time pressure* (6.18) involved [6.17:6.18]. In general, we would expect selective perception to be most acute where time is shortest. The relations among these variables are indicated in Figure 6.1.

Other cognitive aspects of the definition of the situation. All the statements of the last section apply, *mutatis mutandis*, to the other elements of the definition of the situation besides goals and values. That is to say, the definition of the situation represents a simplified, screened, and biased model of the objective situation, and filtering affects all of the "givens" that enter into the decision process: knowledge or assumptions about future events; knowledge of sets of alter-

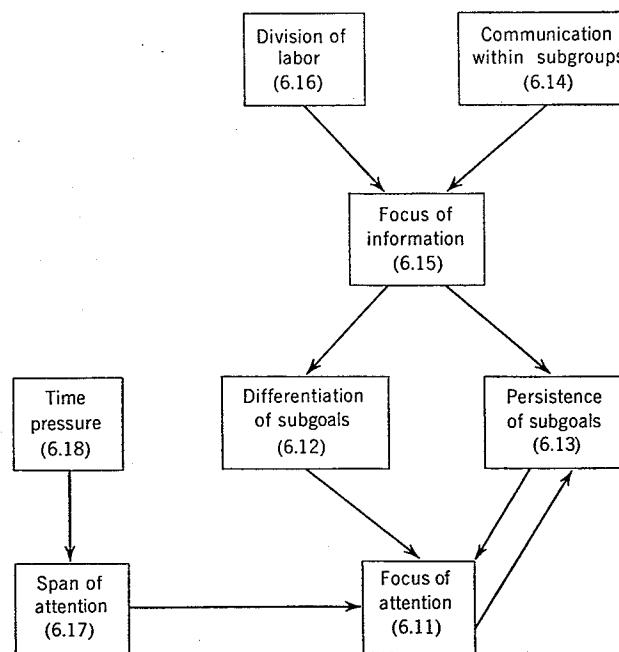


Figure 6.1. Some factors affecting selective attention to subgoals.

natives available for action; knowledge of consequences attached to alternatives; goals and values (Levin, 1956; Gore, 1956).

Consider just knowledge and assumptions about future and present events—"stipulated facts," "absorption of uncertainty." What the sales of the ABC Company are going to be in 1961 is a question of fact. But this matter of fact may become a matter of organizational stipulation—all action within the organization to which the 1961 sales figure is relevant being based upon an "official" sales forecast. Organizational techniques for dealing with uncertain future and present facts will be discussed in a later section of this chapter.

A related phenomenon is the summarizing of raw information to communicate it further in the organization. The weatherman makes observations of temperature, humidity, barometric pressure, but may communicate only his conclusions in the form of a weather forecast. In organizational communication evidence is replaced with conclusions drawn from that evidence, and these conclusions then become the "facts" on which the rest of the organization acts [A-6.15]. One particular form of summarization is classification. When a particular thing has been classified as belonging to a species, all the attributes of the species can be ascribed to the individual instance of it. Priority systems are an example of an important kind of formal classification device.

Similarly, individuals and organizations develop repertoires of programs of action suited to different situations. These are frequently combined with classification systems so that once a situation has been assigned to a particular class the appropriate action program can be applied to it. Such repertoires of performance programs, and the requisite habits and skills for their use, appear to make up the largest part of professional and vocational training.

Knowledge of consequences is intimately related to selective attention to subgoals, and does not require further elaboration here.

The goals that are included in the definition of the situation influence choice only if there are some means, valid or illusory, for determining the connections between alternative actions and goal satisfaction—only if it can somehow be determined whether and to what extent these goals will be realized if particular courses of action are chosen. When a means of testing actions is perceived to relate a particular goal or criterion with possible courses of action, the criterion will be called *operational*. Otherwise the criterion will be called non-operational. This distinction has already been made in discussing the effects of organizational reward systems.

For some purposes we will need to make the further distinction be-

tween cases where means-end relations can be evaluated prior to action, and those where they can be evaluated only after the fact. We will call operational goals in the former case *operational ex ante*, in the latter case *operational ex post*.

The goal of "promoting the general welfare" is frequently a part of the definition of the situation in governmental policy-making. It is a nonoperational goal because it does not provide (either *ex ante* or *ex post*) a measuring rod for comparing alternative policies, but can only be related to specific actions through the intervention of subgoals. These subgoals, whose connection with the broader "general welfare" goal is postulated but not testable, become the operational goals in the actual choice situation. (Speaking strictly, whether a goal is operational or nonoperational is not a yes-no question. There are all degrees of "operationality." It will often be convenient, however, to refer simply to the two ends of the continuum.)

An important circumstance causing the substitution of subgoals for more general goals as criteria of decision is that the former are perceived as operational, the latter as nonoperational [A-6.16]. For example, a business firm may understand to some degree how its specific actions affect its share of market, but may understand less surely how its actions affect long-range profits. Then the subgoal of maintaining a particular share of market may become the effective criterion of action—the operational goal.

The distinction between operational and nonoperational goals, combined with the generalization that behavior in organizations is intendedly rational, leads to the consideration of two qualitatively different decision-making processes associated with these two kinds of goals. When a number of persons are participating in a decision-making process, and these individuals have the same operational goals, differences in opinion about the course of action will be resolved by predominately analytic processes, i.e., by the analysis of the expected consequences of courses of action for realization of the shared goals. When either of the postulated conditions is absent from the situation (when goals are not shared, or when the shared goals are not operational and the operational subgoals are not shared), the decision will be reached by predominately bargaining processes. These are, of course, a distinction and prediction made in Chapter 5 and lead to a proposition previously suggested: Rational, analytic processes take precedence over bargaining processes to the extent that the former are feasible. The condition of feasibility is that there be shared operational goals. The proposition, while it has not been much tested, is eminently testable. The goal structure of participants in a

decision-making process can be determined by observation of their interaction or by interviewing or opinion-polling techniques. Their understanding of the means-end connections, and of possible methods for testing these connections, can be ascertained in the same way. It is not difficult to code their actual interaction in such a way as to detect the amount of bargaining.

The distinction between operational and nonoperational goals has been made the basis for the distinction between unitary and federal organization units (Simon, Smithburg, and Thompson, 1950, pp. 268-72). This distinction will be explored in the next chapter.

The distinction between operational and nonoperational goals also serves to explain why a theory of public expenditures has never developed a richness comparable to that of the theory of public revenues. The economic approach to a theory of public expenditures would postulate some kind of "utility" or "welfare" function. A rational expenditure pattern would be one in which the marginal dollar of expenditure in each direction would make an equal marginal contribution to welfare. Although statements of this kind are encountered often enough in the literature of public finance, they are infrequently developed. The reason is that, in the absence of any basis for making the welfare maximization goal operational (because of the absence of an operational common denominator among the subgoals of governmental service), the general statement leads neither to description nor to prescription of behavior (Simon, 1943).

In the literature on organizations, identification with subgoals has generally been attributed to motivation. Hence, in an analysis of conflict among organizational units, the affective aspects of the conflict have been stressed. In the present section, we have seen that cognitive processes are extremely important in producing and reinforcing subgoal identification. Subgoals may replace broader goals as a part of the whole process of replacing a complex reality with a simplified model of reality for purposes of decision and action (Blau, 1955).

What difference does it make whether subgoal identification is motivationally or cognitively produced—whether the attachment to the subgoal has been internalized or is only indirect, through a cognitive link to some other goal? It may make very little or no difference in the short run; indeed, it may be difficult to find evidence from short-run behavior that would distinguish between these mechanisms. But it may make a great deal of difference in the processes for changing identifications. The greater the *dependence of the identification on cognitive links to other goals* (6.19), the greater the *effectiveness of attention-directing stimuli in changing goal emphasis* (6.20)

[6.20:6.19]. By the same token, where identification depends on cognitive links, the invention of new techniques for evaluating the means-ends connections between action alternatives and goals will transform bargaining processes into processes of rational analysis. These hypotheses can be tested empirically.

6.4 THE DIVISION OF WORK

Insofar as tasks are highly programmed, the division of work is a problem of efficient allocation of activities among individuals and among organizational units—a version of the assignment problem already discussed in Chapter 2. However, we need to make two distinctions that tend to be overlooked in the classical theory: First, there is a problem of specialization among individual employees, and a problem of specialization among organizational units. There is no reason to suppose that both sets of problems have the same answers or that the same general principles apply to both. Second, the division of work that is most effective for the performance of relatively programmed tasks need not be the same as that which is most effective for the performance of relatively unprogrammed tasks. In the present discussion, we shall be concerned primarily with programmed tasks; the subject of unprogrammed tasks will be reserved of the next chapter.

The economies of individual specialization arise principally from opportunities for using programs repetitively [A-6.17]. To develop in a person the capacity to carry out a particular program requires an investment in training. In automatic operations, there is an analogous capital investment in machinery capable of carrying out the program. In the case of a computing machine, a substantial part of this investment actually consists of the cost of programming the machine for the particular operations in question. In all of these cases there are economies to be derived, *ceteris paribus*, from assigning the work so as to minimize this investment cost per unit of program execution.

Programs that are built into machines or acquired by humans usually take the form of generalized means—skills or processing capacities that can be used in executing a wide variety of tasks. Typing skill, for example, is a skill of transforming any manuscript into type-written form, and typing occurs as a subprogram in a wide range of programs. Similarly, a drill press is a bundle of capacities for drilling holes; the program can be called into play whenever the fabrication of some product requires holes to be drilled.

This rather obvious point underlies the central problem in specializing highly programmed activities. Consider an organization that performs a large number of tasks, each consisting of the fabrication of a product. If we analyze the fabrication process into subprograms, we find that it becomes economical to arrange the work so that there will be specialized means (machines and trained employees) for performing some of these subprograms. But since a number of these specialties will be required for the manufacture of each product, we create in this way considerable interdependence and need for coordination among them. The greater the *specialization by subprograms* (6.21) (process specialization), the greater the *interdependencies among organizational subunits* (6.22) [6.22:6.21].

Interdependence does not by itself cause difficulty if the pattern of interdependence is stable and fixed. For in this case, each subprogram can be designed to take account of all the other subprograms with which it interacts. Difficulties arise only if program execution rests on contingencies that cannot be predicted perfectly in advance. In this case, coordinating activity is required to secure agreement about the estimates that will be used as the basis for action, or to provide information to each subprogram unit about the relevant activities of the others. Hence, we arrive at the proposition that the more repetitive and predictable the situation, the greater the *tolerance for interdependence* (6.23) [6.23:6.3]. Conversely, the greater the elements of variability and contingency, the greater is the burden of coordinating activities that are specialized by process (MacMahon, Millet, and Ogden, 1941).

Thus, we predict that process specialization will be carried furthest in stable environments, and that under rapidly changing circumstances specialization will be sacrificed to secure greater self-containment of separate programs [A-6.18]. A second prediction is that organizations, in order to permit a greater degree of process specialization, will devise means for increasing stability and predictability of the environment [A-6.19].

Three important devices come under this heading. All of these devices may be regarded as instances of the more general practice of standardization—of reducing the infinite number of things in the world, potential and actual—to a moderate number of well-defined varieties. The greater the *standardization of the situation* (6.24), the greater the tolerance for subunit interdependencies [6.23:6.24].

The first step in almost all major manufacturing sequences that lead from natural raw materials to finished goods is refining. In steel manufacture, a complex of natural materials—ores, coke, and flux—is

reduced to a relatively homogeneous, standard material—pig iron. In the natural textile industries, fibers are transformed into threads of uniform size, strength, and elasticity by carding and spinning processes. In all such cases, the complexity of subsequent manufacturing processes and their contingency on raw materials is reduced by transforming highly variable natural materials into much more homogeneous semimanufactured products [A-6.20]. After homogeneity has been attained, subsequent steps in the manufacturing process may again produce great variety in the product—alloy steels in the first example, dyed fabrics in the second. But it is often difficult and expensive to program this subsequent elaboration unless the processing begins with a simple, homogeneous material of known properties.

A second important device for dealing with the interdependencies created by specialization is the use of interchangeable parts [A-6.21]. When the fit of two parts is assured by setting minimum and maximum size limits, the interdependency between the units that make them is decreased and the burden of coordination partly removed.

Third, the need for coordinated timing between successive process steps is reduced by holding buffer inventories [A-6.22]. If process A precedes process B in the manufacture of some item, then the effect of variations in the rate of process A upon process B can be largely removed by maintaining an inventory of products on which process A has been completed.

Even with such devices, the need for coordination typically remains. The most common device for securing coordination among subprograms where there is a high degree of process specialization is scheduling. A schedule is simply a plan, established in advance, that determines what tasks will be handled and when. It may have greater or less detail, greater or less precision. The *type of coordination* (6.25) used in the organization is a function of the extent to which the situation is standardized [6.25:6.24]. To the extent that contingencies arise, not anticipated in the schedule, coordination requires communication to give notice of deviations from planned or predicted conditions, or to give instructions for changes in activity to adjust to these deviations. We may label coordination based on pre-established schedules *coordination by plan*, and coordination that involves transmission of new information *coordination by feedback*. The more stable and predictable the situation, the greater the reliance on coordination by plan; the more variable and unpredictable the situation, the greater the reliance on coordination by feedback.

Insofar as coordination is programmed and the range of situations sufficiently circumscribed, we would not expect any particularly close

relation between the coordinative mechanisms and the formal organizational hierarchy. That is to say, scheduling information and feedback information required for coordination are not usually communicated through hierarchical channels. Hierarchy may be important in establishing and legitimizing programs, but the communication involved in the execution of highly programmed activities does not generally follow the "lines of command" [A-6.23] (Bakke, 1950).

In addition, from the standpoint of any particular organization, specialization and the structure of subprograms is as much sociological as it is technological. The organization depends to a great extent upon the training that employees bring to it—training acquired by apprenticeship or in schools. Hence the boundaries of specialization of individual jobs tend to be determined by the structure of trades and professions in the broader social environment [A-6.24].

6.5 COMMUNICATION

On the basis of the foregoing analysis, we may classify the occasions for communication as follows:

1. Communication for nonprogrammed activity. This is a catchall category that will need further analysis later.
2. Communication to initiate and establish programs, including day-to-day adjustment or "coordination" of programs.
3. Communication to provide data for application of strategies (i.e., required for the execution of programs).
4. Communication to evoke programs (i.e., communications that serve as "stimuli").
5. Communication to provide information on the results of activities.

The distinction between the first two categories and the last three is the familiar distinction between communication relating to procedural matters and communication relating to substantive content.

Empirical evidence for the distinction among the last three categories was obtained from a study of the use of accounting data by operating departments in manufacturing concerns. It was found that accounting information was used at various executive levels to answer three different kinds of questions: (a) Problem-solving questions: Which course of action is better? This corresponds to our category 3. (b) Attention-directing questions: What problems shall I look into? This corresponds to category 4. (c) Score-card questions:

How well am I (or is he) doing? This corresponds to category 5. Some of the accounting information was also used in connection with less programmed activity (Simon, Guetzkow, Kozmetsky, and Tyndall, 1954). We will consider this point below.

Communication and coordination. The capacity of an organization to maintain a complex, highly interdependent pattern of activity is limited in part by its capacity to handle the communication required for coordination. The greater the *efficiency of communication* (6.26) within the organization, the greater the tolerance for interdependence [6.23:6.26]. The problem has both quantitative and qualitative aspects.

As we noted earlier, it is possible under some conditions to reduce the volume of communication required from day to day by substituting coordination by plan for coordination by feedback. By virtue of this substitution, organizations can tolerate very complex interrelations among their component parts in the performance of repetitive activities. The coordination of parts is incorporated in the program when it is established, and the need for continuing communication is correspondingly reduced. Each specific situation, as it arises, is largely covered by the standard operating procedure.

A different method for increasing the organization's tolerance for interdependence is to increase the efficiency of communication by making it possible to communicate large amounts of information with relatively few symbols. An obvious example is the blueprint, which provides a common plan stated in extreme detail. The blueprint employs a carefully defined, highly developed "language" or set of symbolic and verbal conventions. Because of this standardized language, it can convey large quantities of information. The same attention to standardization of language is seen in accounting systems and other reporting systems that employ numerical data.

Accounting definitions and blueprint conventions are examples of a still more general phenomenon: technical languages, whose symbols have definite and common meanings to the members of an organization. Prominent in these technical languages are categories for classifying situations and events.

The role of unambiguous technical terms in permitting coordination by feedback is shown by the Christie-Luce-Macy experiments (Macy, Christie, and Luce, 1953) with "noisy marbles" in the Bavelas network. Participants in the experiment were given some colored marbles, and they were required to discover what color was held by all of them. Control groups were given marbles that had solid colors like "red," "yellow," etc. Experimental groups were given streaked

marbles whose colorings did not correspond in any simple way to color designations in common language. Comparison of the performance of the control with the experimental groups showed (a) that the latter were much hindered by the lack of adequate technical vocabulary, and (b) that their performance became comparable to that of the control groups only when they succeeded in inventing such a vocabulary and securing its acceptance throughout the group.

Classification schemes are of particular significance for the program-evoking aspects of communication. When an event occurs that calls for some kind of organization response, the question is asked, in one form or other: "What kind of event is this?" The organization has available a repertory of programs, so that once the event has been classified the appropriate program can be executed without further ado. We can make this process more specific with a pair of examples.

The oil gauge on the dashboard of an automobile is an example of the use of classification in program-evoking. For most drivers, the oil pressure is either "all right" or "low." In the first case, no action is taken; in the second case a remedial program is initiated (e.g., taking the automobile to a repair shop). Some auto manufacturers have substituted a red light, which turns on when the oil pressure is not in the proper range, for the traditional gauge. This example also illustrates how substituting standards of satisfactory performance for criteria of optimization simplifies communication.

Similarly, inspection activities often involve dichotomous decisions. In these cases, the choice is not usually between evoking a program or not evoking one (action or inaction), but between different programs. Thus, if the item being inspected meets the standards, one program is evoked (it is passed on for further processing); if it fails to meet standards, another program is evoked (scrapping, or reworking, as the case may be).

One reason that classifying is so economical of communication is that most of the coordination can be preprogrammed; the organization has a repertory of responses to stimuli, and it only needs to know what kind of stimulus it is confronted with in order to execute an elaborate program. On the other hand, if the communication system could handle a more complete description of the program-evoking event, and if the action part of the organization had the capacity to develop programs on the spot to meet present needs, no doubt one could conceive tailor-made programs that would be more accurately adapted to each separate situation than are the preprogrammed responses.

Here again the normative or adaptive problem of organization de-

sign is one of balance. If its model of reality is not to be so complex as to paralyze it, the organization must develop radical simplifications of its responses. One such simplification is to have (a) a repertory of standard responses, (b) a classification of program-evoking situations, (c) a set of rules to determine what is the appropriate response for each class of situations. The balance of economies and efficiencies here is exactly the same as it is in all cases of standardization. Note that what we have described in an organizational framework is quite comparable to discrimination learning in individuals. In the individual case, as in the organizational, there is a close relationship between the categories used in the cognitive code and the operational decision rules (Whorf, 1956).

In our culture, language is well developed for describing and communicating about concrete objects. The blueprint has already been mentioned as an important technical device for this purpose. Language is also very effective in communicating about things that can be classified and named, even if they are intangible. Thus, when there are standard repertoires of programs, it is easy to refer to them.

On the other hand, it is extremely difficult to communicate about intangible objects and nonstandardized objects. Hence, the heaviest burdens are placed on the communications system by the less structured aspects of the organization's tasks, particularly by activity directed toward the explanation of problems that are not yet well defined. We shall see in the next chapter that this difference in communication difficulty has important implications for the organization of nonprogrammed activities.

Where the available means of communication are primitive—relative to the communication needs—so will be the system of coordination. There will tend to be less self-containment of organizational units and a greater reliance on coordination through communication the greater the efficiency of communication [6.12:6.26]. This relation may sometimes be obscured by the fact that pressure toward coordination (e.g., under conditions of rapid change) may compel attempts at feedback coordination even though available communication is inefficient. It should also be noted that self-containment decreases and interdependencies increase the likelihood of developing an efficient communication code [6.26:6.21].

The absorption of uncertainty. The use of classification schemes in communication has further consequences, some of which go back to our earlier discussion of perception and identification. The technical vocabulary and classification schemes in an organization provide a set of concepts that can be used in analyzing and in communicating

about its problems. Anything that is easily described and discussed in terms of these concepts can be communicated readily in the organization; anything that does not fit the system of concepts is communicated only with difficulty. Hence, the world tends to be perceived by the organization members in terms of the particular concepts that are reflected in the organization's vocabulary. The particular categories and schemes of classification it employs are reified, and become, for members of the organization, attributes of the world rather than mere conventions (Blau, 1955).

The reification of the organization's conceptual scheme is particularly noticeable in *uncertainty absorption* (6.27). Uncertainty absorption takes place when inferences are drawn from a body of evidence and the inferences, instead of the evidence itself, are then communicated. The successive editing steps that transform data obtained from a set of questionnaires into printed statistical tables provide a simple example of uncertainty absorption.

Through the process of uncertainty absorption, the recipient of a communication is severely limited in his ability to judge its correctness. Although there may be various tests of apparent validity, internal consistency, and consistency with other communications, the recipient must, by and large, repose his confidence in the editing process that has taken place, and, if he accepts the communication at all, accept it pretty much as it stands. To the extent that he can interpret it, his interpretation must be based primarily on his confidence in the source and his knowledge of the biases to which the source is subject, rather than on a direct examination of the evidence.

By virtue of specialization, most information enters an organization at highly specific points. Direct perception of production processes is limited largely to employees in a particular operation on the production floor. Direct perception of customer attitudes is limited largely to salesmen. Direct evidence of the performance of personnel is restricted largely to immediate supervisors, colleagues, and subordinates.

In all of these cases, the person who summarizes and assesses his own direct perceptions and transmits them to the rest of the organization becomes an important source of informational premises for organizational action. The "facts" he communicates can be disbelieved, but they can only rarely be checked. Hence, by the very nature and limits of the communication system, a great deal of discretion and influence is exercised by those persons who are in direct contact with some part of the "reality" that is of concern to the organization. Both the amount and the *locus of uncertainty absorption* (6.28) affect the *influence structure of the organization* (6.29) [6.29:6.27, 6.28].

Because of this, uncertainty absorption is frequently used, consciously and unconsciously, as a technique for acquiring and exercising power. In a culture where direct contradiction of assertions of fact is not approved, an individual who is willing to make assertions, particularly about matters that do not contradict the direct perceptions of others, can frequently get these assertions accepted as premises of decision.

We can cite a number of more or less "obvious" variables that affect the absorption of uncertainty. The more complex the data that are perceived and the less adequate the organization's language, the closer to the source of the information will the uncertainty absorption take place, and the greater will be the amount of summarizing at each step of transmission. The locus of absorption will tend to be a function of such variables as: (a) the needs of the recipient for raw as against summarized information (depending upon the kinds of data used in selecting the appropriate program), (b) the need for correction of biases in the transmitter, (c) the distribution of technical competence for interpreting and summarizing raw data, and (d) the need for comparing data from two or more sources in order to interpret it.

The way in which uncertainty is absorbed has important consequences for coordination among organizational units. In business organizations, expected sales are relevant to decisions in many parts of the organization: purchasing decisions, production decisions, investment decisions, and many others. But if each organizational unit were permitted to make its own forecast of sales, there might be a wide range of such estimates with consequent inconsistencies among the decisions made by different departments—the purchasing department, for example, buying raw materials that the production department does not expect to process. It may be important in cases of this kind to make an *official* forecast and to use this *official* forecast as the basis for action throughout the organization.

Where it is important that all parts of an organization act on the same premises, and where different individuals may draw different conclusions from the raw evidence, a formal uncertainty absorption point will be established, and the inferences drawn at that point will have official status in the organization as "legitimate" estimates. The greater the need for coordination in the organization, the greater the *use of legitimized "facts"* (6.30) [6.30:6.8, 6.9].

The communication network. Associated with each program is a set of information flows that communicate the stimuli and data required to evoke and execute the program. Generally this communica-

tion traverses definite channels, either by formal plan or by the gradual development of informal programs. Information and stimuli move from sources to points of decision; instructions move from points of decision to points of action; information of results moves from points of action to points of decision and control.

Rational organization design would call for the arrangement of these channels so as to minimize the communication burden. But insofar as the points of origin of information and the points of action are determined in advance, the only mobile element is the point of decision. Whatever may be the position in the organization holding the formal authority to legitimize the decision, to a considerable extent the effective discretion is exercised at the points of uncertainty absorption.

In large organizations, specialization of communication functions will be reflected in the division of work itself. Among the specialized communication units we find are (a) units specializing in the actual physical transmission of communications: a telephone and teletype unit, messenger group, or the like; (b) units specializing in recording and report preparation: bookkeeping and other record-keeping units; (c) units specializing in the acquisition of raw information, usually referred to as intelligence units, sometimes as research units; (d) units specializing in the provision of technical premises for decision: research units, technical specialists; (e) units specializing in the interpretation of policy and organizational goals, a function usually not much separated from the main stem of the hierarchy; and (f) units specializing in the retention of information: files, archives units [A-6.25].

In part, communication channels are deliberately and consciously planned in the course of programming. In part, they develop through usage. We will make two hypotheses about such development. First, the greater the communication efficiency of the channel, the greater the *communication channel usage* (6.31) [6.31:6.26]. The possession by two persons, or two organization units, of a common, efficient language facilitates communication. Thus, links between members of a common profession tend to be used in the communication system. Similarly, other determinants of language compatibility—ethnic background, education, age, experience—will affect what channels are used in the organization.

Second, channel usage tends to be self-reinforcing [6.31:6.31]. When a channel is frequently used for one purpose, its use for other unrelated purposes is encouraged. In particular, formal hierarchical channels tend to become general-purpose channels to be used when-

ever no special-purpose channel or informal channel exists or is known to the communicator. The self-reinforcing character of channel usage is particularly strong if it brings individuals into face-to-face contact. In this case (the Homans hypothesis) informal communication, much of it social in character, develops side-by-side with task-oriented formal communication, and the use of the channel for either kind of communication tends to reinforce its use for the other.

In part, the communication network is planned; in part, it grows up in response to the need for specific kinds of communication; in part, it develops in response to the social functions of communication. At any given stage in its development, its gradual change is much influenced by the pattern that has already become established. Hence, although the structure of the network will be considerably influenced by the structure of the organization's task, it will not be completely determined by the latter.

Once a pattern of communication channels has become established, this pattern will have an important influence on decision-making processes, and particularly upon nonprogrammed activity. We may anticipate some of the analysis of the next chapter by indicating briefly the nature of this influence.

The existing pattern of communication will determine the relative frequency with which particular members of the organization will encounter particular stimuli, or kinds of stimuli, in their search processes [6.11:6.31]. For example, a research and development unit that has frequent communication with sales engineers and infrequent communication with persons engaged in fundamental research will live in a different environment of new product ideas than a research and development unit that has the opposite communication pattern.

The communication pattern will determine how frequently and forcefully particular consequences of action are brought to the attention of the actor. The degree of specialization, for example, between design engineers, on the one hand, and installation and service engineers, on the other, will have an important influence on the amount of awareness of the former as to the effectiveness of their designs.

From our previous propositions concerning time pressure effects, we would predict that the pattern of communication would have a greater influence on nonprogrammed activities carried out with deadlines and under time pressure than upon activities that involve relatively slow and deliberate processes of decision. For, given sufficient time, if particular information is available anywhere in an organization, its relevance to any particular decision is likely to be noticed.

Where decisions are made relatively rapidly, however, only the information that is locally available is likely to be brought to bear. We see here another reason why specialization (in this case specialization with respect to possession of information) is tolerated to a greater degree under "steady-state" conditions than when the organization is adapting to a rapidly changing environment.

6.6 ORGANIZATION STRUCTURE AND THE BOUNDARIES OF RATIONALITY

It has been the central theme of this chapter that the basic features of organization structure and function derive from the characteristics of human problem-solving processes and rational human choice. Because of the limits of human intellective capacities in comparison with the complexities of the problems that individuals and organizations face, rational behavior calls for simplified models that capture the main features of a problem without capturing all its complexities.

The simplifications have a number of characteristic features: (1) Optimizing is replaced by satisficing—the requirement that satisfactory levels of the criterion variables be attained. (2) Alternatives of action and consequences of action are discovered sequentially through search processes. (3) Repertoires of action programs are developed by organizations and individuals, and these serve as the alternatives of choice in recurrent situations. (4) Each specific action program deals with a restricted range of situations and a restricted range of consequences. (5) Each action program is capable of being executed in semi-independence of the others—they are only loosely coupled together [A-6.26].

Action is goal-oriented and adaptive. But because of its approximating and fragmented character, only a few elements of the system are adaptive at any one time; the remainder are, at least in the short run, "givens." So, for example, an individual or organization may attend to improving a particular program, or to selecting an appropriate program from the existing repertory to meet a particular situation. Seldom can both be attended to simultaneously.

The notion that rational behavior deals with a few components at a time was first developed extensively in connection with economic behavior by John R. Commons, who spoke of "limiting factors" that become the foci of attention and adaptation. Commons' theory was further developed by Chester I. Barnard, who preferred the term "strategic factor."

This "one-thing-at-a-time" or "*ceteris paribus*" approach to adaptive

behavior is fundamental to the very existence of something we can call "organization structure." Organization structure consists simply of those aspects of the pattern of behavior in the organization that are relatively stable and that change only slowly. If behavior in organizations is "intendedly rational," we will expect aspects of the behavior to be relatively stable that either (a) represent adaptations to relatively stable elements in the environment, or (b) are the learning programs that govern the process of adaptation.

An organization is confronted with a problem like that of Archimedes: in order for an organization to behave adaptively, it needs some stable regulations and procedures that it can employ in carrying out its adaptive practices. Thus, at any given time an organization's programs for performing its tasks are part of its structure, but the least stable part. Slightly more stable are the switching rules that determine when it will apply one program, and when another. Still more stable are the procedures it uses for developing, elaborating, instituting, and revising programs.

The matter may be stated differently. If an organization has a repertory of programs, then it is adaptive in the short run insofar as it has procedures for selecting from this repertory a program appropriate to each specific situation that arises. The process used to select an appropriate program is the "fulcrum" on which short-run adaptiveness rests. If, now, the organization has processes for adding to its repertory of programs or for modifying programs in the repertory, these processes become still more basic fulcra for accomplishing longer-run adaptiveness. Short-run adaptiveness corresponds to what we ordinarily call problem-solving, long-run adaptiveness to learning.

There is no reason, of course, why this hierarchy of mechanisms should have only three levels—or any specified number. In fact, the adaptive mechanisms need not be arranged hierarchically. Mechanism A may include mechanism B within its domain of action, and vice versa. However, in general there is much asymmetry in the ordering, so that certain elements in the process that do not often become strategic factors (the "boundaries of rationality") form the stable core of the organization structure.

We can now see the relation between Commons' and Barnard's theories of the "limiting" or "strategic" factor and organization structure. Organization will have structure, as we have defined the term here, insofar as there are boundaries of rationality—insofar as there are elements of the situation that must be or are in fact taken as givens, and that do not enter into rational calculations as potential strategic factors. If there were not boundaries to rationality, or if the

boundaries varied in a rapid and unpredictable manner, there could be no stable organization structure. Some aspects of structure will be more easily modified than others, and hence we may need to distinguish short-run and long-run structure.

In this chapter, we have been concerned mostly with short-run structure—with programs to respond to sequences of situations requiring adaptive action. The "boundaries of rationality" that have been the source of our propositions have consisted primarily of the properties of human beings as organisms capable of evoking and executing relatively well-defined programs but able to handle programs only of limited complexity.

In the next chapter, we will shift our attention to long-run considerations, and particularly to the processes in organizations that bring programs into existence and that modify them.